Models and Algorithms for Election Control through Influence Maximization*

Mohammad Abouei Mehrizi, Federico Corò, Emilio Cruciani, Gianlorenzo D'Angelo, and Stefano Ponziani

Gran Sasso Science Institute (GSSI), L'Aquila, Italy
{mohammad.aboueimehrizi, federico.coro, emilio.cruciani, gianlorenzo.dangelo, stefano.ponziani}@gssi.it

Abstract. We present our ongoing work on the election control problem via social influence. We consider the problem of exploiting social influence in a network of voters to change their opinion about a target candidate with the aim of increasing his chance to win or lose the election.
We introduce the *Linear Threshold Ranking* and the *Probabilistic Linear Threshold Rankings*, natural and powerful extensions of the well-established *Linear Threshold Model*. In both models we are able to maximize the score of a target candidate by showing submodularity. We exploit such property to provide a constant factor approximation algorithm for the *constructive* and *destructive* election control problems.
We outline some further research directions which we are investigating.

Keywords: Computational Social Choice · Election Control · Influence Maximization · Social Networks · Voting Systems

1 Introduction

Recently, there has been a growing interest in the relationship between social networks and political campaigning. Political campaigns nowadays use social networks to lead elections in their favor, e.g., by spreading fake news on the elections outcome [13]. There exists evidence of political intervention which shows the effect of social media manipulation. A real-life example is in the recent presidential US election where a study on the effect of social media on people shows that in average 92% percent of American remembered the pro-Trump fake news, and 23% percent of them remembered the pro-Clinton fake news [1]. Many other real-life examples have been recorded and studied [3, 8, 10, 14].

There exists an extensive literature on manipulating elections without considering the underlying social network structure of the voters; we point the reader to a recent survey [7]. Nevertheless, there are only few studies that exploit opinion diffusion in social networks to change the outcome of elections. Independent

* This work has been partially supported by the Italian MIUR PRIN 2017 Project ALGADIMAR "Algorithms, Games, and Digital Markets".

Cascade Model (ICM) [9] has been considered as diffusion process in the problem of constructive/destructive election control [15], that consists in changing voters' opinions with the aim of maximizing/minimizing the margin of victory of target candidate w.r.t. its most voted opponent. A variant of the Linear Threshold Model (LTM) [9] with weights on the vertices has been considered on a graph in which each node represents a cluster of voters with a specific list of candidates and there is an edge between two nodes if they differ by the ordering of a single pair of adjacent candidates [6].

In this work we focus on the election control problem via social influence, introduced in [15]: Given a social network of voters, we want to select a subset of voters that, with their influence, will change the opinion of the network's users about a target candidate, maximizing its chances to win/lose. As in [15], we consider the scenario in which only the opinions about a single candidate can be changed, simulating the spread of a single news among the users of the network. With this aim we start by defining a model in which the preference list of each voter is known [4, 5]. However, this assumption is not always satisfied in a realistic scenario as voters can be undecided on their preferences or they may not reveal them to the manipulator. Thus, we extend this model to design a scenario in which the manipulator can only guess a probability distribution over the candidates for each voter [11]. We consider LTM as diffusion model to take into account the degree of influence that voters exercise on the others to describe the scenario in which a high influence on someone can radically change its opinion. Wilder et al. [15] studied a scenario in which a voter can only change the position of the target candidate in its ranking by one position, regardless of the amount of influence received, using ICM under plurality rule.

In this paper we report our ongoing research on which we wish to get feedback so as to possibly include these results in future publications. We point interested reader to our preliminary published results in [11, 4, 5].

2 Background

Voting Systems are sets of rules that regulate all aspects of elections and determine their outcome. Herein we consider two *single-winner* voting systems: (i) *Plurality rule*: Each voter can only express a single preference among the candidates; the winner is the candidate with the highest number of votes. (ii) *Scoring rule*: Each voter expresses his preference as a *ranking*; each candidate is then assigned a *score*, computed as a function of its position in the ranking. This is a general definition that include several election methods by choosing an adequate *scoring function*: (a) *plurality rule*: 1 point to the first candidate and 0 to the others; (b) t*-approval*: 1 point to the first t candidates and 0 to the others (each voter approves t candidates); (c) *borda count*: $m - l$ points to a candidate in position l, with m the number of candidates.

Influence Maximization is the problem of finding a subset of nodes in a graph that maximizes the spread of information. In this work we focus on the diffusion model known as *Linear Threshold Model* (LTM) [9]. Given a graph $G = (V, E)$, in LTM each node $v \in V$ has a threshold $t_v \in [0, 1]$ sampled independently and uniformly at random and each edge $(u, v) \in E$ has a weight $b_{uv} \in [0, 1]$ such that $\sum_{(u,v) \in E} b_{uv} \leq 1$. Each node can be either *active*, that is it spreads the information, or *inactive*. Let $A_t \subseteq V$ be the set of *active* nodes at time t, where A_0 is the set of active nodes at the beginning of the process. A node v becomes active if the sum of the weights of the incoming edges from active nodes is greater than or equal to its threshold, i.e., $v \in A_t$ if and only if $v \in A_{t-1}$ or $\sum_{u \in A_{t-1} : (u,v) \in E} b_{uv} \geq t_v$. We define the eventual set of active nodes as $A := A_{\tilde{t}}$, where $\tilde{t}$ is the first time in which the set of active nodes does not change anymore, i.e., $A_{\tilde{t}} = A_{\tilde{t}+1}$. Given a budget B, the influence maximization problem consists in finding a set of nodes A_0 of size at most B such that $\mathbf{E}\left[|A|\right]$ is maximum.

A central result is that the distribution of A, starting from any set A_0, is equivalent to the distribution of reachable nodes in the set of random subgraphs called *live-edge graphs* [9]. In a live-edge graph each node v has at most one incoming edge: Each incoming edge (u, v) is sampled with probability proportional to its weight b_{uv} and no incoming edge is sampled with probability $1 - \sum_{u \in V : (u,v) \in E} b_{uv}$. The result states that the influence maximization problem in live-edge graphs is monotone and *submodular*;[1] hence, it can be approximated to a factor of $1 - 1/e$ using a simple hill-climbing greedy algorithm [12]. While it is $\#P$-hard to compute the expected number of active nodes, there exists a simulation-based approach in which the spread of influence can be evaluated by sampling a polynomial number of live-edges [9, Proposition 4.1].

3 Models and Algorithms for Election Control

We present two models for the election control problem, both a variant of the Linear Threshold Model. At the end of LTM, we modify the lists of preference of voters according to some predefined rules. We represent the underlying social network as a directed graph $G = (V, E)$. Let $C = \{c_1, \dots, c_m\}$ be the set of m candidates; we refer to our *target candidate*, i.e., the one that we want to make win/lose the elections, as $c_\star$. Each voter $v \in V$ has a list of preferences for the elections represented as a function $\pi_v : C \to \mathbb{R}$. Then, we define the initial expected *score* of a candidate c_i as the number of votes that c_i obtains from the voters before the process starts, that is, $F(c_i, \emptyset) := \sum_{v \in V} \pi_v(c_i)$, and the *expected score* of a candidate c_i at the end of the process as $F(c_i, S) := \sum_{v \in V} \tilde{\pi}_v(c_i)$, where S is the set of seed nodes.

In the problem of *election control* we want to maximize the chances of the target candidate to win/lose the elections. To achieve that, we maximize its expected *Margin of Victory* (MoV). As in [15], we define $\text{MoV}(S)$ as the expected increase of the difference between the score of $c_\star$ and that of its most voted opponent. Formally, if c and $\hat{c}$ are respectively the candidates different from $c_\star$ with the highest score before and after the LTM process, the MoV is

$$\text{MoV}(S) := F(c, \emptyset) - F(c_\star, \emptyset) - \left(F(\hat{c}, S) - F(c_\star, S)\right).$$

[1] For a ground set N, a function $z : 2^N \to \mathbb{R}$ is *submodular* if for any S, T s.t. $S \subseteq T \subseteq N$ and $\forall e \in N \setminus T$ it holds that $z(S \cup \{e\}) - z(S) \geq z(T \cup \{e\}) - z(T)$.

Given a budget B, the *election control problem* asks to find an initial set of seed nodes S, of size at most B, that maximizes the MoV, i.e.,

$$\arg\max_S \text{MoV}(S) \qquad \text{s.t. } |S| \leq B.$$

3.1 Linear Threshold Ranking

We first introduce a deterministic model called *Linear Threshold Ranking* (*LTR*), based on LTM, that takes into account the degree of influence that voters exercise on each other. As in LTM, each node $v \in V$ has a threshold $t_v \in [0, 1]$; each edge $(u, v) \in E$ has a weight b_{uv} with the constraint that $\sum_{u:(u,v)\in E} b_{uv} \leq 1$. Moreover, each node v has a permutation π_v of C, i.e., its list of preferences for the elections; in this case $\pi_v(c_i)$ denote the position of candidate c_i in the preference list of node v. After the LTM process stops, i.e., no more nodes are being activated, the position of $c_\star$ in the preference list of each node changes according to a function of its incoming active neighbors. Let A be the set of active nodes at the end of LTM. The threshold t_v of each node $v \in V$ models its strength in retaining its original opinion about candidate $c_\star$: The higher is the threshold the lower is the probability that v is influenced by its neighbors. The weight of an edge b_{uv} measures the influence that node u has on node v. The new position of $c_\star$ after the process will be

$$\tilde{\pi}_v(c_\star) := \pi_v(c_\star) - \min\left(\pi_v(c_\star) - 1, \left\lfloor \frac{\alpha(\pi_v(c_\star))}{t_v} \sum_{u\in A,\,(u,v)\in E} b_{uv} \right\rfloor\right),$$

where $\alpha : \{1, \ldots, m\} \to [0, 1]$ is a function that depends on the position of $c_\star$ in π_v and models the rate at which $c_\star$ shifts up; the candidates overtaken by $c_\star$ will shift one position down. In this general model we are able to approximate the maximum score that $c_\star$ can achieve up to a factor of $1 - 1/e$ through the use of Greedy, a greedy algorithm that iteratively selects the node that maximizes the increment in score [12]. We exploit this fact to provide the following results.

Theorem 1. Greedy *gives a* $\frac{1}{3}(1 - 1/e)$*-approximation for the* constructive *and a* $\frac{1}{2}(1 - 1/e)$*-approximation for the* destructive *election control problems in arbitrary scoring rule voting systems, under the LTR model.*

3.2 Probabilistic Linear Threshold Rankings

In this section we extend the previous model and consider a non-deterministic scenario in which we take into account the inherent uncertainty of a voter and we model its decision as a probabilistic function over the list of candidates. Each node $v \in V$ has a probability distribution over the candidates π_v; here $\pi_v(c_i)$ denotes the probability that v votes for candidate c_i; then for each $v \in V$ we have that $\pi_v(c_i) \geq 0$ for each candidate c_i and $\sum_{i=1}^{m} \pi_v(c_i) = 1$. We focus on the *plurality* voting rule. Given an initial set of seed nodes S, the diffusion process proceeds as in LTM; then, at the end of the process, active nodes increase their probability of voting for the target candidate by adding the influence coming from the active neighbors and then by normalizing to have again a probability distribution. Formally, for each node $v \in A$, where A is the set of active nodes at the end of LTM, the preference list π_v changes as follows:

$$\tilde{\pi}_v(c_\star) = \frac{\pi_v(c_\star) + \sum_{u \in A \cap N_v^i} b_{uv}}{1 + \sum_{u \in A \cap N_v^i} b_{uv}}, \quad \text{and} \quad \tilde{\pi}_v(c_i) = \frac{\pi_v(c_i)}{1 + \sum_{u \in A \cap N_v^i} b_{uv}}, \forall c_i \neq c_\star, \quad (1)$$

where N_v^i denote the sets of incoming neighbors for a node $v \in V$. All inactive nodes $v \in V \setminus A$ will have $\tilde{\pi}_v(c_i) = \pi_v(c_i)$ for all candidates, including $c_\star$. We call this process *Probabilistic Linear Threshold Ranking* (PLTR).

For this model, we show that the election control problem is hard to approximate to within a polynomial fraction of the optimum through a reduction from Densest-k-Subgraph problem, hard to approximate within any constant factor. However, we are able to show that a small relaxation of the model allows us to give a constant factor approximation algorithm. We call this model *Relaxed* PLTR (R-PLTR), here, the probability distribution of a node is updated if it has at least an active incoming neighbor, also if the node is not active itself: Every node $v \in V$ updates its probability distribution according to (1) and not just every node $v \in A$ as in PLTR. The rationale is that a voter might slightly change its opinion about the target candidate if it receives some influence from its active incoming neighbors even if the received influence is not enough to activate it (thus making it propagate the information to its outgoing neighbors). Therefore, we include this small amount of influence in the objective function. The election control problem in R-PLTR is still *NP*-hard; however, adapting Greedy for the weighted-LTM case, we are able to guarantee a constant factor approximation ratio in this setting for both the constructive and destructive scenarios.

Theorem 2. Greedy *gives a* $\frac{1}{6}(1 - 1/e)$*-approximation for the* constructive *and a* $\frac{1}{4}(1 - 1/e)$*-approximation for the* destructive *election control problems in the plurality rule voting system, under the R-PLTR model.*

4 Future Research

We presented our ongoing work on the problem of controlling elections through social influence: Given a social network of voters, we want to select a subset of them that, influencing the others about a specific candidate, will make him win/lose the elections. We introduced *Linear Threshold Ranking* and the *Probabilistic Linear Threshold Rankings*, models that describe the change of opinions taking into account the amount of exercised influence, and two greedy algorithms with constant factor approximation to the problems of *constructive* and *destructive* election control in the model. As future research directions we would like to study our model in scenarios which are not currently captured, including *multi-winner* and *proportional representation* systems. It is also worth to analyze approaches that mix constructive and destructive control. Finally, it would be interesting to study how to prevent election control for the integrity of voting processes, e.g., through the placement of monitors in the network [2, 18] or by considering strategic settings [16, 17].

References

1. Allcott, H., Gentzkow, M.: Social media and fake news in the 2016 election. Working Paper 23089, National Bureau of Economic Research (January 2017). https://doi.org/10.3386/w23089, http://www.nber.org/papers/w23089
2. Amoruso, M., Anello, D., Auletta, V., Ferraioli, D.: Contrasting the spread of misinformation in online social networks. In: Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2017, São Paulo, Brazil, May 8-12, 2017. pp. 1323–1331 (2017), http://dl.acm.org/citation.cfm?id=3091308
3. Bond, R.M., Fariss, C.J., Jones, J.J., Kramer, A.D.I., Marlow, C., Settle, J.E., Fowler, J.H.: A 61-million-person experiment in social influence and political mobilization. Nature **489**(7415), 295–298 (2012). https://doi.org/10.1038/nature11421, https://doi.org/10.1038/nature11421
4. Corò, F., Cruciani, E., D'Angelo, G., Ponziani, S.: Exploiting social influence to control elections based on scoring rules. In: Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19. pp. 201–207. International Joint Conferences on Artificial Intelligence Organization (7 2019). https://doi.org/10.24963/ijcai.2019/29, https://doi.org/10.24963/ijcai.2019/29
5. Corò, F., Cruciani, E., D'Angelo, G., Ponziani, S.: Vote for me!: Election control via social influence in arbitrary scoring rule voting systems. In: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems. pp. 1895–1897. AAMAS '19, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2019), http://dl.acm.org/citation.cfm?id=3306127.3331955
6. Faliszewski, P., Gonen, R., Koutecký, M., Talmon, N.: Opinion diffusion and campaigning on society graphs. In: Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden. pp. 219–225 (2018). https://doi.org/10.24963/ijcai.2018/30, https://doi.org/10.24963/ijcai.2018/30
7. Faliszewski, P., Rothe, J.: Control and bribery in voting. In: Handbook of Computational Social Choice, pp. 146–168. Cambridge University Press (2016). https://doi.org/10.1017/CBO9781107446984.008, https://doi.org/10.1017/CBO9781107446984.008
8. Ferrara, E.: Disinformation and social bot operations in the run up to the 2017 french presidential election. First Monday **22**(8) (2017), http://firstmonday.org/ojs/index.php/fm/article/view/8005
9. Kempe, D., Kleinberg, J.M., Tardos, É.: Maximizing the spread of influence through a social network. Theory of Computing **11**, 105–147 (2015). https://doi.org/10.4086/toc.2015.v011a004, https://doi.org/10.4086/toc.2015.v011a004
10. Kreiss, D.: Seizing the moment: The presidential campaigns use of twitter during the 2012 electoral cycle. New Media & Society **18**(8), 1473–1490 (2016)
11. Mehrizi, M.A., Corò, F., Cruciani, E., D'Angelo, G.: Election control with voters' uncertainty: Hardness and approximation results. CoRR **abs/1905.04694** (2019), http://arxiv.org/abs/1905.04694
12. Nemhauser, G.L., Wolsey, L.A., Fisher, M.L.: An analysis of approximations for maximizing submodular set functions - I. Math. Program. **14**(1), 265–294 (1978). https://doi.org/10.1007/BF01588971, https://doi.org/10.1007/BF01588971
13. Pennycook, G., Cannon, T., Rand, D.G.: Prior exposure increases perceived accuracy of fake news. Journal of Experimental Psychology: General (2018)

14. Stier, S., Bleier, A., Lietz, H., Strohmaier, M.: Election campaigning on social media: Politicians, audiences, and the mediation of political communication on facebook and twitter. Political Communication **35**(1), 50–74 (2018). https://doi.org/10.1080/10584609.2017.1334728
15. Wilder, B., Vorobeychik, Y.: Controlling elections through social influence. In: Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2018, Stockholm, Sweden, July 10-15, 2018. pp. 265–273 (2018), http://dl.acm.org/citation.cfm?id=3237428
16. Wilder, B., Vorobeychik, Y.: Defending elections against malicious spread of misinformation. CoRR **abs/1809.05521** (2018), http://arxiv.org/abs/1809.05521
17. Yin, Y., Vorobeychik, Y., An, B., Hazon, N.: Optimal defense against election control by deleting voter groups. Artif. Intell. **259**, 32–51 (2018). https://doi.org/10.1016/j.artint.2018.02.001, https://doi.org/10.1016/j.artint.2018.02.001
18. Zhang, H., Alim, M.A., Thai, M.T., Nguyen, H.T.: Monitor placement to timely detect misinformation in online social networks. In: 2015 IEEE International Conference on Communications, ICC 2015, London, United Kingdom, June 8-12, 2015. pp. 1152–1157 (2015). https://doi.org/10.1109/ICC.2015.7248478, https://doi.org/10.1109/ICC.2015.7248478